Using Diaries for Evaluating Interactive Products: The Relevance of Form and Context

Martin Tomitsch, Nikash Singh, Ghazaleh Javadian
Faculty of Architecture, Design and Planning
The University of Sydney 2006 NSW Australia
martin.tomitsch@sydney.edu.au, nikash.singh@sydney.edu.au, ghjavadian@gmail.com

ABSTRACT
In this paper we discuss two studies, in which we used incident diaries to evaluate different aspects of a web-based tool and a wearable display. For the web-based tool we used a diary in form of a table distributed in digital form, which resulted in a very low number of responses. Results from follow-up interviews revealed that one of the reasons for this low response rate was a mismatch between diary form and study context. For the wearable display we designed booklets, which featured predefined sections and questions as well as space for open comments. Although previous research has identified disadvantages of paper-based diaries, this method proved to be valuable for collecting feedback in a mobile context. Based on our experiences and the results from the studies, we provide a qualitative discussion of design issues for diaries used in mobile and desktop-based contexts.

Author Keywords
Diaries, evaluation studies, mobile, wearable, web.

ACM Classification Keywords
H5.2. Information interfaces and presentation (e.g., HCI): User Interfaces – Evaluation/methodology.

INTRODUCTION

Diaries are a well-established method of recording self-reported data from users. Compared to other methods, such as surveys or interviews, the usage of diaries has the advantage that users record events when they occur (Carter and Mankoff, 2005). This reduces the risk of participants providing unreliable data, since humans are often not able to accurately recall past events (Schacter, 1999). On the other hand, filling in diaries represents a distraction from the current task, which can affect the results recorded in diaries (Czerwinski et al., 2004).

In human-computer interaction (HCI) research, there are two types of studies in which diaries are used – feedback studies and elicitation studies (Carter and Mankoff, 2005). In feedback studies participants are prompted with questions and asked to record information about an event immediately after it occurred. In elicitation studies participants only capture aspects of the event when it occurs (such as photographs), which is used by the researcher later in interviews to prompt the participant's memory to reveal further information about the event. Feedback studies therefore lead to more accurate results, since the information is recorded in situ, but can also be more burdensome depending on the number of questions and the frequency of events (Carter and Mankoff, 2005). Incident diaries are a form of diary used in feedback studies. They are similar in format to questionnaires and are particularly useful in studies with a low number of events (Baxter and Oatley, 1991). Most feedback studies rely on paper-based diaries, although it was noted that this can potentially have an impact on the sustained participation of participants (Carter and Mankoff, 2005). Ways to address this include periodic reminders and reimbursement strategies (Palen and Salzman, 2002).

In HCI research, diaries are typically either used to understand how people complete daily activities in order to direct the design of new technologies (e.g. Czerwinski et al., 2004; Palen and Salzman, 2002) or to evaluate usability and user experience aspects of novel interactive products (e.g. Lichtner et al., 2009). Using diaries in evaluation studies is especially useful in situations where it is not possible to automatically log events or user behaviour while participants are using the product due to technical or privacy-related issues.

While diaries have been used in many HCI studies and the method is well covered in HCI literature, only little information is available that helps researchers to choose the right diary format for the context of their study. Carter and Mankoff (2005) have addressed this lack of research by evaluating the role of media in elicitation studies. This paper intends to extend their findings regarding feedback studies, where researchers are not faced with the question of which media to capture, but which form of diary to use. To discuss this, we draw on two studies, in which we used a combination of paper-based and digital incident diaries.

RELATED WORK

Palen and Salzman (2002) carried out a paper diary study to find out how people learn to use computer systems in the workplace and outside of an educational environment. Participants were asked to fill in log sheets about their computer-related activities during their working hours. Sellen and Harper (1997) provide another example of early use of a feedback paper diary study in an HCI context, where they sought to investigate work practices in a document-intensive environment and understand how the relevant technology can better assist workers' task processes. In a similar work-related design exercise,

OZCHI 2010, November 22-26, 2010, Brisbane, Australia.

OZCHI 2010 Proceedings ISBN: 978-1-4503-0502-0

Adler et al. (1998) used a diary study to provide insight to the design of digital reading devices at work. Other feedback studies have investigated non-paper based forms of diaries. For example, Palen and Salzman (2002) have introduced the voice-mail diary study technique to examine mobile telephony use. In their study, participants were asked to record activities by calling a specific voice mail line. Czerwinski et al. (2004) used a simple digital diary where participants recorded task events in a spreadsheet, with the goal to uncover new software design opportunities to assist everyday work patterns. Carter and Mankoff (2005) described another approach of using mobile phones for data collection. They used a combination of feedback and elicitation studies to explore how people make public transit decisions. Participants were asked to make a phone call every time they made a transit decision and answer a series of questions, whilst location was captured from built in GPS sensors to provide information for later elicitation. Since in this study, events (transit decisions) occurred less frequently, the common drawback of feedback studies imposing a burden on participants was alleviated. The results consequently showed a relatively high response rate of feedback from participants. The evolution of diary studies within HCI is further highlighted by the more recent study of Lichtner et al. (2009) utilising online forums as a medium for diary reporting in the evaluation of a work-integrated learning system. They also included questions in the forum to prompt users, however the majority of users preferred to record experiences in their own words. In total the diary study resulted in 87 entries from 24 users over a period of 4 weeks. This high number of responses can most likely be attributed to the fact that the diary was designed for the context of use – both, the system of inquiry as well as the diary, were running on the users' desktop computer. Lichtner et al. also proposed a set of design principles for designing digital online diaries. These guidelines emphasised the importance of factors such as flexibility, usability and interactivity for a successful diary tool. A similar set of principles were earlier put forward by Palen and Salzman (2002) in the context of voicemail diary reporting, some of which could also be considered more widely by any diary method.

DIARY STUDIES

This section provides an outline of the studies referred to in this paper. In particular, the section gives a brief overview of study methodologies and describes the diaries used in the studies in detail. Both studies were carried out at our research group over the course of the previous year. In both studies we used incident diaries to evaluate a novel interactive product. In the first study we used a diary to record events while participants were using a browser add-on that allowed them to quickly connect to friends in social online networks. In the second study we used a paper-based diary to evaluate a wearable display.

Browser Add-on Study

This study investigated the impact of embedding a person's closest friends within their web browser and integrating social channels such as Facebook and Flickr within this interaction. We developed a browser add-on, which provided one-click access to five close friends and six online channels including email. The aim was to observe whether users felt more connected and aware of their friends. We provided participants with a diary to record whether feelings of social connection and increased awareness were actually experienced. Participants used the diary during the one-week long deployment study while using the browser add-on on their own computers. The diary was then used again to document a week of online social interactions that occurred without the use of the browser add-on. Ten participants took part in the study (4 female, 6 male) ranging in age from 20 to 28 with occupations including information workers and students.

We asked participants to record information after every online social interaction as well as at the end of the day. The diary had the form of a table, with each horizontal record constituting one online social interaction. Participants were asked to specify which friends they contacted and which social channels they used. Each interaction also required a rating for the urgency of the communication and a rating for recovery time. At the end of every day, participants had to provide a rating for how aware they felt of each of their five friends and for how inconvenienced they felt by each of the social channels used that day.

We distributed the diary as Microsoft Word file to all participants via email as some of them were based in another country. Additionally we handed out a paper printout of the diary to the 5 participants we had direct access to when we met them for a briefing session. These 5 participants returned the diary on paper filled in by hand, the other 5 remote participants returned the Word file filled in digitally.

There were several problems regarding the design of the diary, partly caused by an attempt to condense the diary into a single page to make it convenient for participants to keep the diary handy. For example the daily ratings table was at the top of the page while participants expected this to be at the bottom. Having to provide a rating for every friend and every social channel used that day was also taxing on participant's time and memory. Many felt obligated to provide ratings despite not having contacted a specific friend or having used a particular channel that day.

In the follow-up interviews, one participant noted "The unpleasant aspects came from having to log my activities everyday, which made each action tedious, as opposed to the joy of spontaneously using the program." while a second participant mentioned "I don't know I guess filling any diary can be a bit annoying, but I guess it makes what you do a more conscious action as you are doing it. And it can be hard to rate how urgent or whatever something is on a scale, because often when I do it I'm not thinking about it." All participants unanimously agreed that they would have preferred an online diary. Reasons identified for this preference ranged from the comparable ease of commenting by typing, through to the convenience of an integrated solution to compliment the digital nature of the

prototype being evaluated. One participant stated "Since it is a web-based program, an online form of diary might be more convenient and it can be plugged into your program as well".

Wearable Display Study

The aim of this study was to evaluate a wearable visualisation and sonification display, which represents live sensor-based information (Beilharz et al., 2010). This information is visible to the wearers and people in their immediate vicinity. The intention of this wearable display is to augment social interaction and interplay. Information is represented through physical (folds in fabric) as well as auditory means with an emphasis on the aesthetic integration of the display to support wearable expression.

Five participants (3 female, 2 male; age range 26-40; 4 students, one employee) took part in the evaluation study. All participants were selected based on their demographics and background, working in an environment that would allow wearing the display for most part of the day and being interested in new technologies and gadgets. Participants were asked to wear the display for three days during their normal daily activities.

We used a diary to allow participants to capture their responses in situ. The diary was iteratively refined and continuously tested in several pilot test sessions. The diary was 14 by 9.5cm (5.5 by 3.7 inches) in size to fit into trouser pockets and handbags. This ergonomic design was crucial to allow participants to capture events in various circumstances and contexts. The cover of the diary booklet was kept in a plain, unobtrusive and neutral colour. For privacy reasons and to ensure the participant would feel comfortable writing into it in any environment no label was included. The first page of the diary featured a welcome message and basic instructions (Figure 1, top). In this study, there were two different events that participants had to capture: *onlooker events* and *device events*. For each event we included a double page populated with questions to prompt responses. This included fields such as time, location, and current activity as well as information about the onlookers, such as their behaviour, and any interaction they had with the onlookers. Emoticons were included to identify the onlookers' reaction to the device. For the device event (i.e. for changes in the physical display or occurring sounds), we asked participants to capture the devices' behaviour and their interpretation of this behaviour. For both events additional space was provided at the end of each double page to record other comments or to draw sketches (Figure 1, bottom). The two types of events were repeated in alternating order throughout the booklet. At the end of the booklet we included a timeline capturing the entire evaluation period, where participants could indicate whether they wore the display and provide reasons.

At the end of the 3-day evaluation period we conducted interviews with each participant. In these interviews the diaries were used to prompt the participants' memory. For example, one of the participants stated that there was no

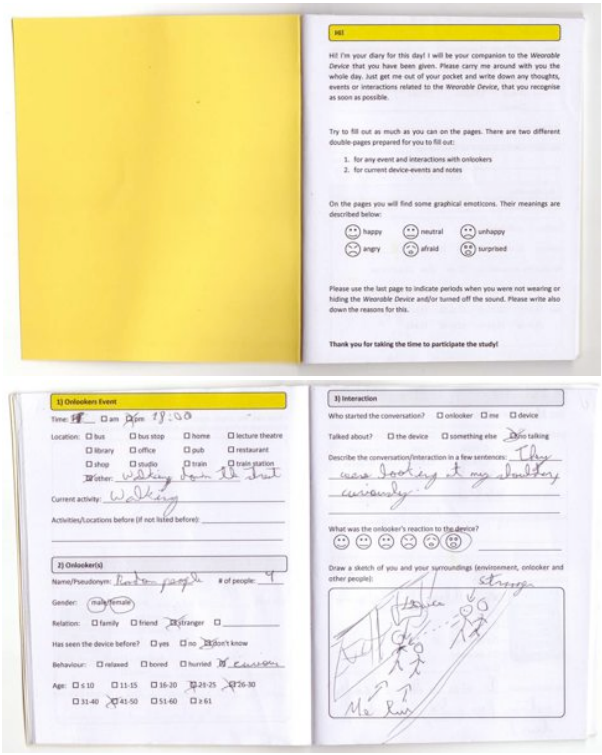

Figure 1. The first page of the diary used in the wearable display study, showing the welcome message (top), and the layout design of the diary (filled in by one of the participants), which allows for recording notes as well as sketches (bottom).

stranger looking at the display on a specific day. After the researcher pointed out the captured event in her diary, she recalled the event and was able to provide further details.

DISCUSSION

Many of the issues why participants in the browser add-on study failed to record events correctly were due to usability flaws and a lack of instructions. It is important that participants receive instructions on how to use their diary (Carter and Mankoff, 2005). In the wearable display study we included these instructions on the first page of the diary, so participants were able to always look them up during the study. Similarly, online diaries and other digital diaries should provide instructions. Rather than trying to compress all the fields that need to be recorded into one single page, as we did in the browser add-on study, enough room needs to be provided to contain these instructions.

All participants from the browser add-on study expressed in follow-up interviews that they would have preferred an online diary. We believe that one of the reasons for this is the context, in which the study was carried out. The add-on was designed to be used while browsing the Internet on a desktop computer. We assumed that providing the option to fill in the diary either digitally or on paper would allow participants to record events in a form of their preference without being burdensome or disruptive. This approach was similar to a study, where participants received an Excel spreadsheet to record task switching and interruptions (Czerwinski et al., 2004). However, while in this study using an Excel spreadsheet might have

been an appropriate form, as participants were recording information while using their desktop computer, in our study participants should have been allowed to stay within their environment, the browser. Neither the paper nor the Microsoft Word version of the diary seemed to fulfill this requirement as both led to equally unsuccessful results.

Another reason why using a paper-based form of diary was more successful in the wearable display study can be attributed to the design of the diary. We experienced that participants perceived the diary as a gift, since it was aesthetically designed and looked like a handcrafted journal, adding a very personal character to the object. The welcome message on the first page further added to this character. In contrast, the table-based diary, which participants in the add-on study received, looked very plain and boring.

Regarding the time period of the study, participants from the wearable display study expressed that they liked filling in the diaries during the three-days long study. Although one of the participants explicitly stated that she felt relaxed after the study ended, since she did not have to write into her diary any longer. What is interesting is that she was the only participant in this study, who captured an increasing number of events from day to day. There were no concerns regarding the frequency of events or the amount of details required to be filled in for each event. This might have been due to the fact that the basic information about an event could be filled out really quickly. The other fields were optional and participants could decide themselves whether they wanted to provide this additional information.

Diaries can suffer the disadvantage of possibly missing entries forgotten by the participants to record or not valued to be relevant for the research (Sohn et al., 2008). However, a comparison of the diary entries from the wearable display study with log entries of device events showed exact matches between the description of the perceived events by the wearers and actual occurrence of events. In fact, the results from the wearable display study suggest that the booklet acted as unobtrusive physical reminder, which triggered participants to record events when they occurred. This feature would be lost when replacing the paper-based diary with a digital form of diaries.

In some mobile contexts, the use of voice mail diaries as proposed by Palen and Salzman (2002) can be beneficial for the participants. However, this approach does not allow participants to provide sketches for events, which was important in our wearable display study and has also been identified as a relevant requirement in previous research (Carter and Mankoff, 2005).

CONCLUSIONS

The results from our studies indicate that paper-based diaries work well for evaluating interactive products in a mobile context. It is important however to carefully design the aesthetics and the usability of the diary. Advantages of paper-based diaries that we identified in our research include the fact that they act as physical reminders, participants can wear them on their body at all times, and they can record sketches for events. Depending on the context and information that needs to be recorded, other forms should be considered as well. For example, in screen-based products, certain aspects, such as questions regarding the usefulness of an application, can also be directly embedded in the mobile interface. For a successful diary study it is crucial to carefully reflect on the context of the study and in which participants will be using the diary to record events. For example, participants in our browser add-on study failed to record events, and stated they would have preferred using online diaries, which would have allowed them to stay within the browser environment when recording events.

ACKNOWLEDGMENTS

We thank Andrew Vande Moere, who was involved in designing both studies, and Barbara Stiel, who was involved in the wearable display study.

REFERENCES

Adler, A., Gujar, A., Harrison, B.L., O'Hara, K and Sellen, A. A diary study of work-related reading: design implications for digital reading devices. In Proc. CHI '98, ACM Press (1998), 241-248

Baxter, I. and Oatley, K. Measuring the learnability of spreadsheets in inexperienced users and those with previous spreadsheet experience. In Behaviour & Information Technology, 10, 6 (1991), 475-490.

Beilharz, K., Vande Moere, A., Stiel, B., Calvo, C., Tomitsch, M. and Lombard, A. Expressive Wearable Sonification and Visualisation: Design and Evaluation of a Flexible Display. In Proc. NIME 2010, 323-326.

Carter, S. and Mankoff, J. When participants do the capturing: the role of media in diary Studies. In Proc. CHI '05, ACM Press (2005), 899-908.

Czerwinski, M., Horvitz, E. and Wilhite, S. A diary study of task switching and interruptions. In Proc. CHI '04, ACM Press (2004), 175-182.

Rieman, J. The diary study: A workplace-oriented research tool to guide laboratory efforts. In Proc. INTERACT and CHI '93, ACM Press (1993), 321-326.

Lichtner, V., Kounkou, A. P., Dotan, A., Kooken, J. P. and Maiden, N. M. A. An online forum as a user diary for remote workplace evaluation of a work-integrated learning system. In Proc. CHI '09, ACM Press (2009), 2955-2970.

Palen, L. and Salzman, M. Voice-mail diary studies for naturalistic data capture under mobile conditions. In Proc. CSCW '02, ACM Press (2002), 87-95.

Schacter, D. L. The seven sins of memory: Insights from psychology and cognitive neuroscience. American Psychology, 54 (1999), 182-203.

Sellen, A. and Harper, R. Paper as an analytic Resource for the design of new technologies. In Proc. CHI '97, ACM Press (1997), 319-326.

Sohn, T., Li, K. A., Griswold, W. G. and Hollan, J. D. A diary study of mobile information needs. In Proc. CHI '08, ACM Press (2008), 433-442.